*Page: 241-279* *URL: www.lijdlr.com*

LAWFOYER INTERNATIONAL JOURNAL OF DOCTRINAL LEGAL RESEARCH

ISSN: 2583-7753

# LawFoyer International Journal of Doctrinal Legal Research

**[ISSN: 2583-7753]**

**Volume 4 | Issue 4**

**2026**

*DOI: https://doi.org/10.70183/lijdlr.2026.v04.386*

*© 2026 LawFoyer International Journal of Doctrinal Legal Research*

*Follow this and additional research works at: www.lijdlr.com*
*Under the Platform of LawFoyer – www.lawfoyer.in*

*After careful consideration, the editorial board of LawFoyer International Journal of Doctrinal Legal Research has decided to publish this submission as part of the publication.*

*In case of any suggestions or complaints, kindly contact (info.lijdlr@gmail.com)*
*To submit your Manuscript for Publication in the LawFoyer International Journal of Doctrinal Legal Research, To submit your Manuscript Click here*

# MANDATORY LABELLING AND PROVENANCE METADATA FOR AI-GENERATED CONTENT IN INDIA: A CONSTITUTIONAL EVALUATION OF THE IT RULES, 2021, AS AMENDED IN 2026

**Prakhar Singh**[1]

## I. ABSTRACT

*India's Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, in force from 20 February 2026, require intermediaries that enable synthetic audio-visual content to label lawful synthetically generated information prominently, to embed permanent metadata or other provenance mechanisms carrying a unique identifier to the extent technically feasible, and to prevent the removal of both. Significant social media intermediaries must collect user declarations and verify them before publication. This paper asks whether these obligations are constitutionally justified, sufficiently precise and technically proportionate. Using doctrinal, analytical and comparative methods, it reads the Gazette text against sections 79 and 87 of the Information Technology Act, 2000, Articles 14, 19 and 21 of the Constitution and the structured proportionality test, and draws on technical and empirical studies of provenance and labels. It compares Article 50 of the EU AI Act, China's 2025 labelling measures and California's AI Transparency Act. The paper finds that a narrowly defined, factual disclosure of synthetic origin can survive scrutiny as a modest burden that serves the audience's interest in knowing what it sees. The framework is vulnerable at four points: the absence of any accommodation for satire and evident fiction, the unqualified pre-publication verification duty, undefined technical terms such as "computer resource" and "permanent" metadata, and the lack of notice and appeal for users whose content is labelled in error. The paper proposes a differentiated model of context-sensitive labels,*

[1] LL.M (Cyber law), 1st year, 1st semester, Student at Uttar Pradesh State Institute Of Forensic Science, Lucknow, India. Email: prakharsingh0745@gmail.com.

© 2026. LawFoyer International Journal of Doctrinal Legal Research (ISSN: 2583-7753)

*privacy-protective identifiers, published technical standards and procedural remedies that serves the regulatory objective at lower constitutional cost.*

## II. KEYWORDS

Synthetically Generated Information, Mandatory Labelling, Provenance Metadata, Constitutional Proportionality, Intermediary Liability.

## III. INTRODUCTION

On 10 February 2026, the Ministry of Electronics and Information Technology ("MeitY") notified the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026 ("Amendment Rules"), which came into force on 20 February 2026.[2] The Amendment Rules insert into the Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 ("IT Rules") a definition of "synthetically generated information" ("SGI") and a dedicated set of due diligence duties around it.[3]

Intermediaries that offer tools capable of creating or spreading SGI must prevent certain unlawful synthetic content, label all other SGI, embed provenance metadata carrying a unique identifier and refuse to let anyone strip the label or the metadata. Significant social media intermediaries ("SSMIs") must, before publication, ask users to declare whether their content is synthetic and verify that declaration through technical measures.[4]

The problem these rules address is real. Realistic synthetic audio and video now allow impersonation, fraud, sexual exploitation and political deception at a cost and scale that earlier manipulation techniques never reached. Chesney and Citron add a subtler harm: as fabrications multiply, people caught on genuine recordings can more

[2] The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, r. 1(2), G.S.R. 120(E), Gazette of India, Extraordinary, pt. II sec. 3(i) (Feb. 10, 2026) (India).
[3] The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, r. 2(1) (wa), G.S.R. 139(E), Gazette of India, Extraordinary, pt. II sec. 3(i) (Feb. 25, 2021) (India).
[4] Amendment Rules, 2026, r. 4(a).

© 2026. LawFoyer International Journal of Doctrinal Legal Research (ISSN: 2583-7753)

plausibly dismiss them as fakes, a phenomenon they call the "liar's dividend".[5] A rule that helps audiences tell captured media from generated media therefore serves interests in individual dignity as well as in the quality of public discourse.

Yet the rules reach well beyond harmful content. Rule 3(3)(a)(ii) attaches a label and metadata to every item of SGI that is not otherwise prohibited, which includes lawful art, satire, advertising, teaching material and political commentary that happens to look real. The duties fall formally on intermediaries, but the expression that carries the label belongs to users, and the intermediary's fear of losing its statutory immunity under section 79 of the Information Technology Act, 2000 ("IT Act") will shape how far it labels, verifies and removes.[6] The constitutional question is therefore not whether the State may act against deepfakes, which it plainly may, but whether this particular design is a valid exercise of delegated power and a proportionate limitation of fundamental rights.

This paper answers that question as the law stood on 25 October 2026. After tracing the regime's evolution, research design and literature (Parts 2 to 4), it explains the technical concepts (Part 5), examines the rules and their statutory basis (Parts 6 and 7), applies Articles 14, 19 and 21 (Part 8), compares three foreign models (Part 9) and states its findings and recommendations (Parts 10 to 12).

### A. Research Problem, Questions, Objectives and Methodology

#### 1. Statement of the Research Problem

The research problem is whether the labelling and provenance obligations in the amended IT Rules are constitutionally justified, sufficiently precise, technically proportionate and consistent with fundamental rights. Two considerations pull in opposite directions. Labels and provenance signals may help audiences identify

[5] Robert Chesney & Danielle Keats Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. 1753 (2019).
[6] The Information Technology Act, 2000, § 79, No. 21, Acts of Parliament, 2000 (India); *see also* Chinmayi Arun, *Gatekeeper Liability and Article 19(1)(a) of the Constitution of India*, 7 NUJS L. REV. 73 (2014).

synthetic material, deter deceptive impersonation and support accountability. Compulsory labels, persistent identifiers and pre-publication verification may also burden lawful expression, expose information about creators, impose costs that fall hardest on smaller firms and produce errors that mark authentic content as fake. The paper treats the outcome as open and asks whether less restrictive designs could achieve substantially the same objective.

## 2. Research Questions

- What is the scope of the 2026 amendments concerning SGI, mandatory labels and provenance metadata?
- What statutory authority under the IT Act supports these obligations, and how do they relate to the safe harbour in section 79?
- Are mandatory labels and provenance requirements compatible with Article 19(1)(a), read with the grounds in Article 19(2)?
- Do the rules satisfy the requirements of legality, rational connection, necessity, balancing and procedural fairness?
- What implications do the requirements carry for privacy, anonymous expression and the other interests protected by Article 21?
- Are the requirements sufficiently clear, technically feasible and proportionate across different kinds of content, creators and intermediaries, including under Articles 14 and 19(1)(g)?
- What legal, institutional and technical safeguards would improve the framework while preserving its legitimate objectives?

## 3. Research Objectives

The objectives track these questions: to map the framework governing SGI; to analyse the legal and technical operation of labels, metadata and identifiers; to test them against Articles 14, 19 and 21 and the limits of delegated legislation; to identify procedural gaps; and to propose reforms. The literature review sharpened one

objective: the study treats labels, metadata, identifiers and verification as distinct instruments, because existing commentary tends to assess them together.

## 4. Methodology and Scope

The study uses doctrinal, analytical and comparative methods. Its primary materials are the IT Act, the English text of G.S.R. 120(E) as published in the Gazette of India and the principal IT Rules.[7] It reads delegated legislation according to its plain words, in the context of the enabling provisions in sections 79 and 87, and against the grounds on which Indian courts review subordinate legislation. MeitY's Frequently Asked Questions ("FAQs") serve as an aid to understanding executive intent, not as law, as the FAQs themselves state.[8] Constitutional doctrine comes from decisions of the Supreme Court and High Courts, read for their ratio and distinguished from observations not necessary to the decision.

The comparative component examines the European Union, China and California. The European Union offers a horizontal statute that separates machine-readable marking from human-facing disclosure; China operates the closest design analogue to India, combining explicit and implicit labels with platform verification; and California distinguishes system provenance data from personal provenance data, which bears directly on Article 21. Comparison serves to identify design choices, not to import foreign constitutional standards.

The study involves no interviews, surveys or experiments. Technical claims rest on published standards and peer-reviewed studies, and reasoned inferences are marked as such. Several 2026 developments are documented mainly in official notices and specialist reporting, and pinpoint references appear only where the source could be checked.

---

[7] Amendment Rules, 2026.

[8] MINISTRY OF ELEC. & INFO. TECH., FREQUENTLY ASKED QUESTIONS ON THE INFORMATION TECHNOLOGY (INTERMEDIARY GUIDELINES AND DIGITAL MEDIA ETHICS CODE) AMENDMENT RULES, 2026 1 (2026), https://www.meity.gov.in/static/uploads/2025/10/065b6deb585441b5ccdf8be42502a49c.pdf.

### B. Review of Literature and Research Gap

#### 1. Online Speech, Intermediaries and the Indian Constitution

Bhatia's account of Indian free-speech law remains the most sustained doctrinal treatment of Article 19. Working through the case law historically, he argues that the Article 19(2) grounds should be read narrowly and that vague restrictions corrode free speech by delegating the line between lawful and unlawful speech to officials.[9] His analysis predates generative AI, but it supplies the vagueness and overbreadth framework that this paper applies to terms such as "likely to be perceived as indistinguishable".

Arun's work on gatekeeper liability turns to the intermediary. She argues that when the State makes intermediaries liable for what users say, it effectively conscripts private actors into censorship, and that the intermediary's incentive to avoid liability leads it to remove lawful speech.[10] Balkin generalises the point. He describes a "new school" of speech regulation in which the State acts through the owners of digital infrastructure, producing collateral censorship and weakening the procedural protections that attend direct regulation.[11]

Klonick, by contrast, studies how platforms build their own rule systems and suggest that they respond partly to users and norms, not only to state pressure.[12] The disagreement matters: platforms already label AI content under their own policies, so the question is how state compulsion changes the design and incentives of labelling. None of these authors addresses compulsory labelling of lawful synthetic media.

#### 2. Deepfakes, Harm and Legal Responses

Chesney and Citron offer the leading taxonomy of deepfake harms, from extortion and sabotage of individuals to distortion of elections and erosion of trust in evidence.

---

[9] GAUTAM BHATIA, OFFEND, SHOCK, OR DISTURB: FREE SPEECH UNDER THE INDIAN CONSTITUTION (Oxford Univ. Press 2016).
[10] Arun, *supra* note 5.
[11] Jack M. Balkin, *Free Speech Is a Triangle*, 118 COLUM. L. REV. 2011 (2018).
[12] Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598 (2018).

They survey technical, market and legal responses and caution that detection is an arms race that technology alone will not win.[13] Their analysis is framed by American First Amendment law and focuses on liability for harmful fakes. It does not evaluate a regime that requires every realistic synthetic item, harmful or not, to carry a label.

### 3. The Empirical Effects of Labels

Pennycook and co-authors found that attaching warnings to some false headlines increased the perceived accuracy of false headlines that carried no warning, an "implied truth effect".[14] Altay and Gilardi, in two preregistered experiments with 4,976 participants in the United States and the United Kingdom, found that labelling headlines as AI-generated reduced their perceived accuracy and participants' willingness to share them, whether the headlines were true or false and whether humans or AI had written them; participants tended to assume that "AI-generated" meant full automation without human oversight.[15] These studies concern text headlines and Western samples, not Indian audiences or video. They nonetheless establish two propositions that a court assessing proportionality should not ignore: incomplete labelling can lend false credibility to unlabelled fakes, and labels can depress belief in true content.

### 4. Provenance Technology

The technical literature is candid about limits. NIST's 2024 overview of digital content transparency describes provenance tracking, watermarking and detection, and warns that transparency signals can create a false sense of trust, for instance when authentic footage circulates out of context.[16] The Coalition for Content Provenance and Authenticity ("C2PA") specification, the most widely adopted open standard,

---

[13] Chesney & Citron, *supra* note 4.
[14] Gordon Pennycook et al., *The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines Without Warnings*, 66 MGMT. SCI. 4944 (2020).
[15] Sacha Altay & Fabrizio Gilardi, *People Are Skeptical of Headlines Labeled as AI-Generated, Even if True or Human-Made, Because They Assume Full AI Automation*, 3 PNAS NEXUS pgae403 (2024).
[16] NAT'L INST. OF STANDARDS & TECH., NIST AI 100-4, REDUCING RISKS POSED BY SYNTHETIC CONTENT: AN OVERVIEW OF TECHNICAL APPROACHES TO DIGITAL CONTENT TRANSPARENCY (2024), https://doi.org/10.6028/NIST.AI.100-4.

acknowledges that platforms which do not support it routinely strip embedded metadata, and it provides "soft bindings", such as invisible watermarks and fingerprints, to recover provenance records after stripping.[17] This literature evaluates engineering, not law, and does not ask what a legal duty to keep metadata "permanent" can sensibly mean.

### 5. Indian Commentary on the 2026 Rules

Commentary on the Amendment Rules is so far mostly practitioner notes and policy submissions. The Internet Freedom Foundation welcomed the narrowed definition and exclusions but maintained that labels and identifiers raise compelled speech concerns, that pre-publication verification acts as prior restraint, and that the identity-disclosure consequences in Rule 3(1) (ca) lack a requirement of prior judicial order.[18]

A submission from the Law and Technology Research Group at Jawaharlal Nehru University on the draft Second Amendment argued that several provisions exceed the rule-making power in section 87, that the misrepresentation category in Rule 3(3)(a)(i)(IV) lacks a carve-out for satire and mimicry, that the feasibility qualifier applies to metadata but not to the visible label, that "computer resource" is undefined, and that re-uploading through a platform that strips metadata defeats traceability.[19] These advocacy documents identify the right pressure points but do not carry the analysis through each stage of proportionality.

---

[17] COAL. FOR CONTENT PROVENANCE & AUTHENTICITY, C2PA IMPLEMENTATION GUIDANCE, https://c2pa.org/public-draft/Guidance.html. (last visited Oct. 25, 2026); COAL. FOR CONTENT PROVENANCE & AUTHENTICITY, C2PA TECHNICAL SPECIFICATION VERSION 2.2: SOFT BINDING RESOLUTION API (2025), https://c2pa.org/specifications/specifications/2.2/softbinding/Decoupled.html.

[18] Indumugi C. & Anam, *supra* note 11.

[19] Prabhat Katyayan Mishra et al., *Feedback on the Draft Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Second Amendment Rules, 2026*, CTR. FOR THE STUDY OF L. & GOVERNANCE, JAWAHARLAL NEHRU UNIV. (May 7, 2026), https://www.jnu.ac.in/sites/default/files/cslg/Feedback_ITGuidelines_DigitalMediaEthicsCode.pdf.

### 6. Proportionality Scholarship

Chandra's study of Indian proportionality doctrine argues that courts often invoke the test without applying its stages rigorously, particularly the necessity stage, which asks whether a less restrictive measure would serve the aim.[20] Her critique supplies a methodological caution: an assessment of the SGI rules must state the alternatives concretely and test them, not merely recite the four prongs.

### 7. The Research Gap

The gap is not simply that the amendment is new. The literature divides into four strands that do not meet: constitutional scholarship on online speech that predates synthetic media; deepfake scholarship framed by foreign law and focused on harmful content; technical and empirical work that does not engage Indian doctrine; and Indian commentary on the 2026 rules that identifies objections without a structured proportionality analysis. No work located in this study separates the visible label, the audible prefix, embedded metadata, the unique identifier, anti-tampering and pre-publication verification, and tests each against Articles 14, 19 and 21 using the technical and empirical evidence at the rational connection and necessity stages. That integrated, instrument-by-instrument evaluation is the contribution this paper offers.

## IV. BACKGROUND AND EVOLUTION OF SYNTHETIC MEDIA REGULATION IN INDIA

Before 2026, Indian law dealt with deepfakes through general instruments. Victims relied on the offences of identity theft, cheating by personation and privacy violation in the IT Act, on the Bharatiya Nyaya Sanhita, 2023, and on civil actions for defamation and the protection of personality.[21] In *Anil Kapoor v. Simply Life India*, for example, the Delhi High Court restrained the unauthorised commercial exploitation of an actor's

[20] Aparna Chandra, *Proportionality in India: A Bridge to Nowhere?*, 3 U. OXFORD HUM. RTS. HUB J. 55 (2020).
[21] IT Act, 2000, §§ 66C, 66D, 66E; The Bharatiya Nyaya Sanhita, 2023, No. 45, Acts of Parliament, 2023 (India).

name, likeness and voice, including through AI tools.[22] The IT Rules already required intermediaries to tell users not to host content that impersonates another person, but they contained no concept of synthetic media and no labelling duty.[23]

MeitY published draft amendments on 22 October 2025. The draft defined synthetic information broadly, as any information created or altered by a computer resource that reasonably appears to be authentic or true, and industry bodies warned that much of the content on the internet would need a label under it.[24] The draft also proposed a fixed label covering at least ten per cent of a visual display, or the first ten per cent of an audio clip.[25] Civil society submissions objected that the labels amounted to compelled speech, that pre-publication verification operated as prior restraint and that the draft overrated the reliability of detection and watermarking tools.[26] Reports also link the timing of the final rules to a MeitY direction to X in January 2026, issued after users employed its Grok tool to generate obscene imagery.[27]

The notified rules responded in part. They confined SGI to realistic audio-visual material that appears indistinguishable from a real person or event, added three exclusions for routine editing, document preparation and accessibility uses, and dropped the ten per cent threshold in favour of a qualitative standard of prominence.[28] They also sharply shortened takedown and grievance timelines.

The regime has not stood still. On 30 March 2026, MeitY published draft Second Amendment Rules that would make compliance with ministry "clarifications, advisories" and similar instruments part of an intermediary's due diligence under

---

[22] Anil Kapoor v. Simply Life India, CS(COMM) No. 652 of 2023 (Del. H.C. Sept. 20, 2023) (India).
[23] IT Rules, 2021, r. 3(1)(b).
[24] Anjali Tripathi, *Exclusive: Amid Protests, MeitY Widens Labelling for Social Media Beyond Synthetic Content*, MEDIANAMA (Oct. 8, 2026), https://www.medianama.com/2026/10/223-meity-labelling-social-media-synthetic-content/.
[25] *IT Rules 2026*, IPLEADERS BLOG (2026), https://blog.ipleaders.in/it-rules-2026/.
[26] Indumugi C. & Jhanvi Anam, *IT Intermediary Amendment Rules, 2026 Contradict Their Purpose*, INTERNET FREEDOM FOUND. (Feb. 11, 2026), https://internetfreedom.in/it-intermediary-amendment-rules-2026-contradict-their-purpose/.
[27] *2026 Amendments to India's Information Technology Rules*, LAW.ASIA (Oct. 2026), https://law.asia/india-it-rules-2026-ai-synthetic-media-compliance/.
[28] Amendment Rules, 2026, r. 2(a); *IT Rules 2026*, IPLEADERS BLOG, *supra* note 10.

section 79.[29] On 21 April 2026, it added a proposal to amend Rule 3(3)(a)(ii) so that the SGI label would remain continuously and clearly visible throughout the duration of visual content, and extended comments to 7 May 2026.[30] This research found no notification bringing the Second Amendment into force as of 25 October 2026, and the paper therefore treats the continuous-label proposal as draft, not law. On 8 October 2026, MeitY also issued an advisory asking social media intermediaries to "label/contextualise manipulated, synthetically generated or misleading content", a request broader than the SGI definition in the Amendment Rules.[31] An advisory is not a rule made under section 87.

Two pieces of litigation frame the constitutional context. In *Kunal Kamra v. Union of India*, the Bombay High Court struck down the 2023 amendment that empowered a government fact check unit to flag content about the "business of the Central Government", holding it violative of Articles 14 and 19 and beyond the IT Act; the Union's appeal is pending before the Supreme Court.[32] Separately, the Supreme Court transferred to the Delhi High Court the petitions challenging the IT Rules themselves, which remain pending.[33] This study found no reported judgment, as of the research date, ruling on the SGI provisions.

---

[29] Ministry of Elec. & Info. Tech., Draft Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Second Amendment Rules, 2026 (Mar. 30, 2026) (India), https://www.meity.gov.in/static/uploads/2026/03/30591fc6e322dcbcc9dae84a0f02e9e7.pdf.

[30] Ministry of Elec. & Info. Tech., Extension of Time for Inviting Feedback/Comments of Stakeholders on the Draft Amendments to the IT Rules, 2021, Including Additional Amendments in Rule 3(3)(a)(ii) (Apr. 21, 2026) (India), https://www.meity.gov.in/static/uploads/2026/04/ec197f1206279efb4964965f0dede6c1.pdf, Rohit Singh, *MeitY Tightens AI Label Rules, Mandates Continuous Disclosure*, MEDIANAMA (Apr. 23, 2026), https://www.medianama.com/2026/04/223-meity-ai-label-rules-mandates-continuous-disclosure/.

[31] Tripathi, *supra* note 9.

[32] Kunal Kamra v. Union of India, W.P. (L) No. 9792 of 2023, 2024 LiveLaw (Bom) 482 (Bom. H.C. Sept. 26, 2024) (India); *Kunal Kamra v. Union of India*, GLOBAL FREEDOM OF EXPRESSION, COLUMBIA UNIV., https://globalfreedomofexpression.columbia.edu/cases/kunal-kamra-v-union-of-india/. (last visited Oct. 25, 2026).

[33] *IT Rules, 2021*, INTERNET FREEDOM FOUND., https://internetfreedom.in/tag/it-rules-2021/. (last visited Oct. 25, 2026).

## V. CONCEPTUAL AND TECHNICAL FOUNDATIONS

Generative AI produces new content from learned patterns; synthetic media is its audio-visual output. Manipulated media is broader and covers any alteration of captured content, while automated enhancement, such as noise reduction, uses algorithms without creating a new depiction. The Indian definition of SGI cuts across these categories. It covers content that is artificially or algorithmically "created, generated, modified or altered", but only where the result appears real and depicts a person or event in a way that is, or is likely to be perceived as, indistinguishable from reality, and it excludes routine edits that do not materially alter meaning.[34] The definition is therefore a test of realism and potential to deceive, not a test of technique.

Seven concepts must be kept apart. A *visible label* is a notice that a human viewer sees. An *audible disclosure* performs the same function for audio. *Embedded metadata* is machine-readable information stored in the file, which human viewers do not see unless software displays it. A *unique identifier* is a value within that metadata that distinguishes one item or one source. *Other provenance mechanisms* include invisible watermarks embedded in the signal itself and fingerprints computed from the content. *Authenticity verification* usually means a cryptographic check that content has not changed since a trusted party signed it. *Attribution* links content to a tool, system or organisation; *identification* links it to a natural person. A provenance system can attribute an image to a particular generator without revealing who prompted it.

The rule text matters on this point. Rule 3(3)(a)(ii) requires an identifier "to identify the computer resource of the intermediary used to create, generate, modify or alter such information".[35] On its face, the identifier points to the intermediary's system, not to the user. Whether it becomes a means of identifying people depends on what else the intermediary records and links to it, a question taken up in Part 8.

---

[34] IT Rules, 2021, r. 2(1)(wa).
[35] *Id.* r. 3(3)(a)(ii).

Three technical limits shape the legal analysis. First, metadata is fragile. Screenshots, re-encoding and platforms that do not support provenance standards routinely remove it, which is why C2PA developed soft bindings to recover detached records.[36] Second, watermarks are more durable but can be degraded, removed or forged, and detection tools produce false positives and false negatives.[37] Third, provenance establishes origin, not truth. A cryptographically signed photograph of a real event can still mislead when presented with a false caption, and an unsigned recording may be entirely genuine. Provenance is evidence about how content was made; it cannot settle whether a statement is accurate. Finally, open-source models run locally involve no intermediary, so rules addressed to intermediaries cannot reach such content at creation.

## VI. THE STATUTORY FRAMEWORK AS AMENDED

### A. Source of Power

The Amendment Rules recite that the Central Government made them under section 87(1) and clauses (z) and (zg) of section 87(2) of the IT Act.[38] Clause (zg) authorises rules on the guidelines to be observed by intermediaries under section 79(2). Section 79(1) exempts an intermediary from liability for third-party information, and section 79(2)(c) conditions that exemption on the intermediary observing "due diligence" and any other guidelines the Central Government prescribes.[39] Rule 7 of the IT Rules supplies the sanction: an intermediary that fails to observe the rules loses the section 79(1) exemption and becomes liable under any applicable law.[40]

### B. Definitions and Exclusions

Rule 2(1)(ca) defines "audio, visual or audio-visual information" to include any image, video, sound recording or similar content, however created. Rule 2(1)(wa) then

[36] C2PA SOFT BINDING API, *supra* note 29.
[37] NIST AI 100-4, *supra* note 28.
[38] Amendment Rules, 2026, pmbl.
[39] IT Act, 2000, §§ 79(1), 79(2)(c), 87(2)(zg).
[40] IT Rules, 2021, r. 7.

defines SGI through four cumulative elements: the content is audio-visual; it is artificially or algorithmically created, generated, modified or altered using a computer resource; it "appears to be real, authentic or true"; and it depicts an individual or event in a manner that is, or is likely to be perceived as, indistinguishable from a natural person or real-world event.[41] A proviso excludes (a) routine or good-faith editing, enhancement or compression that does not materially alter substance, context or meaning; (b) good-faith preparation of documents and educational or research materials that creates no false record; and (c) uses solely for accessibility, translation or discoverability that manipulate no material part.[42]

Three features of the definition deserve note. Text falls outside it entirely, as MeitY confirms.[43] The exclusions protect workflow uses, not genres of expression: there is no exclusion for satire, parody, art or evident fiction. MeitY's own explanation is that lawful satire "may be permitted" if it is "appropriately declared/labelled", which confirms that such content is labelled rather than exempt.[44] And the perception standard in the final limb does not say whose perception counts.

Rule 2(1A) provides that references to "information" used to commit an unlawful act, including in Rules 3(1)(b), 3(1)(d), 4(2) and 4(4), include SGI. Rule 2(1B) declares that removal or disabling of information by an intermediary in compliance with the rules, including through automated tools, "shall not amount to a violation" of the conditions in section 79(2)(a) or (b).[45] Part 7 considers whether a rule can make that declaration.

### C. Rule 3(3): Prevention, Labelling and Provenance

Rule 3(3) applies to any intermediary that "offers a computer resource which may enable, permit, or facilitate" the creation, generation, modification, alteration, publication, transmission, sharing or dissemination of SGI.[46] The words "may" and

[41] *Id.* r. 2(1)(ca), (wa).
[42] *Id.* r. 2(1)(wa) proviso.
[43] MEITY FAQS, *supra* note 20, q. 8.
[44] *Id.* q. 7.
[45] IT Rules, 2021, r. 2(1A), (1B).
[46] *Id.* r. 3(3)(a).

"dissemination" bring in both generation tools and distribution platforms, and MeitY reads the provision that way.[47] The rule sets no size threshold.

Clause (a)(i) obliges such intermediaries to deploy "reasonable and appropriate technical measures", including automated tools, so as not to allow users to create or spread SGI that violates any law, and specifically SGI that contains child sexual abuse material, non-consensual intimate imagery or obscene and related content; creates false documents or electronic records; relates to explosives, arms or ammunition; or "falsely depicts or portrays a natural person or real-world event by misrepresenting, in a manner that is likely to deceive", a person's identity, voice, conduct, action or statement, or an event.[48]

Clause (a)(ii) covers "every such information not covered" by clause (a)(i), that is, lawful SGI. It must be "prominently labelled in a manner that ensures prominent visibility in the visual display that is easily noticeable and adequately perceivable, or, in the case of audio content, through a prominently prefixed audio disclosure", so that it can be immediately identified as SGI. The same content "shall be embedded with a permanent metadata or other appropriate technical provenance mechanisms, to the extent technically feasible, including a unique identifier, to identify the computer resource of the intermediary".[49] No technical standard is prescribed. The feasibility qualifier, by its placement, governs the metadata limb only; the visual label carries no such qualifier, an asymmetry already noted in commentary.[50] Clause (b) forbids the intermediary to "enable the modification, suppression or removal" of the label or metadata.[51]

### D. Rule 3(1) (ca) and (cb): Warnings and Consequences

Rule 3(1)(ca) requires Rule 3(3) intermediaries additionally to warn users that misusing their tools for unlawful SGI may attract punishment and that a

[47] MEITY FAQS, *supra* note 20, q. 20.
[48] IT Rules, 2021, r. 3(3)(a)(i).
[49] *Id.* r. 3(3)(a)(ii).
[50] Mishra et al., *supra* note 31.
[51] IT Rules, 2021, r. 3(3)(b).

contravention of Rule 3(3)(a)(i) may lead to removal, account suspension and, "in accordance with applicable law", identification of the user and disclosure of that identity to a complainant who is a victim.[52] Rule 3(1)(cb) then directs an intermediary that becomes aware of any violation of Rule 3(3) concerning SGI covered by clause (a)(ii) to take "expeditious and appropriate action", including the actions listed in Rule 3(1)(ca)(ii).[53] Read literally, a failure to label lawful SGI could thus lead to identity disclosure, although clause (ca)(ii) frames those consequences as responses to unlawful SGI. The cross-reference is ambiguous and should be clarified.

### E. Rule 4(1A): Declaration and Verification by SSMIs

An SSMI that lets users display, upload or publish information must, before display or publication, (a) require users to declare whether the information is SGI; (b) deploy appropriate technical measures to verify the accuracy of that declaration, having regard to the nature, format and source of the information; and (c) where the declaration or verification confirms SGI, ensure that it is clearly and prominently displayed with an appropriate label.[54] A proviso deems the SSMI to have failed in due diligence where it "knowingly permitted, promoted, or failed to act upon" SGI in contravention of the rules. An Explanation extends its responsibility to taking "reasonable and proportionate technical measures" to verify declarations and to ensure that no SGI is published without a declaration or label.[55] Rule 4(4), which earlier required SSMIs to "endeavour" to deploy automated tools against specified material, now requires them to "deploy" such measures.[56]

### F. Timelines

The Amendment Rules cut the time to act on court orders or reasoned government intimations under Rule 3(1)(d) from thirty-six hours to three, the time to resolve grievances from fifteen days to seven, the time to act on certain removal grievances

[52] *Id.* r. 3(1)(ca).
[53] *Id.* r. 3(1)(cb).
[54] *Id.* r. 4(1A).
[55] *Id.* r. 4(1A) proviso & expl.
[56] Amendment Rules, 2026, r. 4(b).

from seventy-two hours to thirty-six, and the time to remove intimate or impersonating content on complaint under Rule 3(2)(b) from twenty-four hours to two.[57] The two-hour window serves victims of non-consensual imagery, but the three-hour window for government intimations leaves little time for any assessment of legality.[58]

### G. What Is Not Law

Three categories of material must be kept distinct from the operative rules. The FAQs explain MeitY's reading but disclaim legal effect.[59] The draft Second Amendment, including the continuous-label proposal and Rule 3(4), remains a draft.[60] The 8 October 2026 advisory is an executive communication; it refers to the IT Rules but does not amend them, and it lists no consequences for non-compliance.[61]

## VII. DELEGATED LEGISLATION, SAFE HARBOUR AND ENFORCEMENT AUTHORITY

Subordinate legislation may be challenged for lack of legislative competence, violation of fundamental rights, failure to conform to the parent statute, repugnance to other laws or manifest arbitrariness, though courts extend it less deference than they extend to statutes only on the last ground.[62] Manifest arbitrariness is also an independent ground under Article 14.[63]

The case for the vires of the labelling and metadata duties is respectable. Section 79(2)(c) expressly contemplates "due diligence" and further prescribed guidelines as conditions of the exemption, and section 87(2) (zg) authorises rules on exactly that subject. Labelling lawful SGI and embedding provenance are measures an

[57] IT Rules, 2021, r. 3(1)(d), 3(2)(a)(i), 3(2)(b).
[58] Indumugi C. & Anam, *supra* note 11.
[59] MEITY FAQs, *supra* note 20, at 1.
[60] Draft Second Amendment Rules, *supra* note 14; MeitY Notice of Apr. 21, 2026, *supra* note 15.
[61] Tripathi, *supra* note 9.
[62] Indian Express Newspapers (Bombay) (P) Ltd. v. Union of India, (1985) 1 S.C.C. 641 (India); State of Tamil Nadu v. P. Krishnamurthy, (2006) 4 S.C.C. 517 (India).
[63] Shayara Bano v. Union of India, (2017) 9 S.C.C. 1 (India).

intermediary take "while discharging its duties", and the formal sanction is the withdrawal of an immunity Parliament made conditional, not a new offence.[64]

The case against is stronger for other provisions. First, section 79 is an exemption. Rules under it may define the conditions of immunity, but the Bombay High Court in *Kunal Kamra* held that a rule conditioning safe harbour on the removal of content flagged by a state unit went beyond the Act, because it effectively created a speech restriction the statute did not authorise.[65] Second, section 79(2)(b) requires that the intermediary not "select or modify" information. The prevention duty in Rule 3(3)(a)(i) requires active selection on the ground that content is "likely to deceive", and Rule 2(1B) declares that compliance does not breach section 79(2). A rule cannot amend the statute by declaration, though a court might read the two harmoniously by treating legally mandated removal as falling outside "selection". Third, *Shreya Singhal* read down section 79(3)(b) so that intermediaries act on court orders or government notifications rather than judging the legality of content themselves.[66] A prevention duty keyed to likely deception asks intermediaries to make exactly that judgment, in advance and at scale.

Two qualifications temper this critique. *Shreya Singhal* concerned takedown on private notice; labelling lawful content does not remove it, so its reasoning applies with less force to Rule 3(3)(a)(ii). And in *X Corp. v. Union of India* the Karnataka High Court upheld the use of section 79(3)(b) through the Sahyog portal and held that a foreign company cannot invoke Article 19.[67] That decision, which X Corp announced it would appeal, suggests that the strongest challengers to the SGI rules will be Indian users, creators and companies rather than foreign platforms.

The draft Rule 3(4) would deepen the delegation concern by making compliance with unpublished ministerial advisories a condition of safe harbour.[68] The October 2026

---

[64] IT Act, 2000, §§ 79(2)(c), 87(2)(zg).
[65] Kunal Kamra, 2024 LiveLaw (Bom) 482.
[66] Shreya Singhal v. Union of India, (2015) 5 S.C.C. 1 (India).
[67] X Corp. v. Union of India, W.P. No. 7405 of 2025 (Kar. H.C. Sept. 24, 2025) (India).
[68] Draft Second Amendment Rules, *supra* note 14; Mishra et al., *supra* note 31.

advisory shows why that matters: it asks platforms to label "misleading" content of any kind, a category the notified rules deliberately left out.[69] If such requests acquired binding force without amendment to the rules, the line between rule-making under section 87 and executive instruction would disappear. On the present law, the assessment is that the labelling and metadata duties probably fall within section 79(2)(c), while the prevention duty in Rule 3(3)(a)(i)(IV) and the deemed-failure proviso in Rule 4(1A) is more exposed to an ultra vires challenge.

## VIII. CONSTITUTIONAL ANALYSIS

### A. Whose Rights Are Engaged

Three sets of interests are in play. Users who create and share synthetic content exercise the freedom of speech and expression.[70] Audiences hold a correlative right to receive information, which the Supreme Court has recognised as part of Article 19(1)(a).[71] Intermediaries carry on a business protected by Article 19(1)(g) and enjoy equal protection under Article 14, though foreign companies, on the Karnataka High Court's reasoning, cannot claim Article 19 rights.[72] The duties are addressed to intermediaries, but Rule 4(1A) reaches users directly by requiring them to declare the nature of their content before publication.

### B. Article 19(1)(a): Labels as Compelled Disclosure

The freedom of speech includes a freedom not to speak. In *Bijoe Emmanuel v. State of Kerala*, the Supreme Court protected students who stood respectfully but declined to sing the national anthem.[73] A mandatory label attaches a statement to the user's expression that the user did not choose, so it engages Article 19(1)(a). Indian law does not, however, treat every compelled statement alike. In *Union of India v. Motion Picture Association*, the Court upheld requirements that cinemas exhibit approved educational

[69] Tripathi, *supra* note 9.
[70] INDIA CONST. art. 19, cl. 1(a).
[71] Sec'y, Ministry of Info. & Broad. v. Cricket Ass'n of Bengal, (1995) 2 S.C.C. 161 (India).
[72] X Corp., W.P. No. 7405 of 2025.
[73] Bijoe Emmanuel v. State of Kerala, (1986) 3 S.C.C. 615 (India).

and news films, reasoning that compelled speech that furthers the informational purpose of free speech does not necessarily infringe it.[74] In *Union of India v. Association for Democratic Reforms*, it held that voters' right to know justified compelled disclosure of candidates' antecedents.[75]

These authorities support a distinction between compelled opinion and compelled fact. The SGI label states a fact about how content was made; it does not require the speaker to endorse a view. American law draws a similar line, permitting "purely factual and uncontroversial" disclosures in commercial speech under *Zauderer*, though *National Institute of Family and Life Advocates v. Becerra* confined that leniency and refused to extend it to non-commercial or contested messages.[76] The comparison does not bind Indian courts, but it shows that factual character lowers a disclosure's constitutional weight without eliminating it.

The label is nonetheless not costless. First, its effect is not neutral. The evidence that "AI-generated" labels depress belief in true and human-made content suggests that a label can change how audiences receive a message, not merely inform them.[77] Second, for satire, parody and art, the label may alter meaning. A satirist or filmmaker whose work depends on a moment of uncertainty must break the frame. Third, the burden varies by design. A small corner notice differs from an overlay that persists through every frame, which the draft continuous-label proposal would require.[78] Embedded metadata, by contrast, does not alter what the audience sees or hears, so its burden on expression is slight; its constitutional weight lies in privacy and feasibility.

The prevention duty in Rule 3(3)(a)(i) stands on a different footing. It does not add information to lawful speech but blocks content before publication. Categories such as child sexual abuse material and non-consensual intimate imagery are unprotected

[74] Union of India v. Motion Picture Ass'n, (1999) 6 S.C.C. 150 (India).
[75] Union of India v. Ass'n for Democratic Reforms, (2002) 5 S.C.C. 294 (India).
[76] Zauderer v. Office of Disciplinary Counsel, 471 U.S. 626 (1985); Nat'l Inst. of Family & Life Advocates v. Becerra, 585 U.S. 755 (2018).
[77] Altay & Gilardi, *supra* note 27.
[78] MeitY Notice of Apr. 21, 2026, *supra* note 15.

by any reading of Article 19(2). Category (IV), false depiction "likely to deceive", is different, because a voice impression or a synthetic sketch of a politician necessarily depicts a natural person and, in a literal sense, misrepresents what that person said.[79] Without an accommodation for evident satire, automated prevention will catch protected expression.

### C. Article 19(2): Locating a Permissible Ground

A Constitution Bench in *Kaushal Kishor v. State of Uttar Pradesh* confirmed that the grounds in Article 19(2) are exhaustive and that no additional restriction may be read in.[80] Preventing deception, misinformation or "AI harm" is not itself a listed ground. The Bombay High Court in *Kunal Kamra* also rejected the premise that the State may act as arbiter of truth in the abstract.[81] The labelling duty must therefore either fall outside the concept of a "restriction" or be connected to a listed ground.

The first route has some support in *Motion Picture Association*, which treated informational compulsion as furthering free speech.[82] But it would be unwise to rest the rules on it, because a label imposed on a speaker's content and enforced through loss of immunity looks like a restriction in substance. The second route is more convincing for the core of the definition. Realistic synthetic depictions of identifiable persons expose them to defamation and to impersonation offences, and fabricated depictions of real-world events, such as riots or attacks, threaten public order and can incite offences. Defamation, public order and incitement to an offence are all listed grounds.[83] The nexus is close where content depicts real people or events and is presented as factual. It weakens where synthetic content depicts an invented person, a fictional scene or a matter no reasonable viewer would take as a record of reality.

There is a further argument in the rules' favour that critics tend to overlook. Without a label, much realistic synthetic content would risk falling within clause (a)(i)(IV) as

---

[79] Mishra et al., *supra* note 31; Indumugi C. & Anam, *supra* note 11.
[80] Kaushal Kishor v. State of Uttar Pradesh, (2023) 4 S.C.C. 1 (India).
[81] Kunal Kamra, 2024 LiveLaw (Bom) 482.
[82] Motion Picture Ass'n, (1999) 6 S.C.C. 150.
[83] INDIA CONST. art. 19, cl. 2.

"likely to deceive" and being blocked. The label is what allows such content to circulate lawfully. In that sense, labelling operates as a less restrictive alternative to prohibition. The constitutional difficulty lies not in labelling as such but in its undifferentiated application to content whose connection with any Article 19(2) ground is remote.

### D. Structured Proportionality

Indian courts test restrictions on fundamental rights through four stages: a legitimate aim, a rational connection between the measure and the aim, necessity in the sense that no less restrictive measure would serve the aim comparably, and a proper balance between benefit and harm.[84] In the Aadhaar judgment, the majority applied a moderate form of the necessity test that asks whether alternatives would realise the aim in a real and substantial manner, rather than demanding the single least restrictive option.[85] *Anuradha Bhasin* applied the same structure to restrictions on internet-based expression and insisted that restrictive orders be published so that they can be challenged.[86] The test applies separately to each instrument.

1. **Legitimate aim:** Protecting individuals from impersonation, sexual exploitation and fraud, and protecting audiences from deception about whether a depiction records reality, are legitimate aims. This stage is satisfied for every instrument.
2. **Rational connection:** A visible label plausibly helps viewers recognise synthetic content, and the audible prefix does the same for audio. The connection is weakened, though not broken, by two facts. Coverage is partial, because content generated locally or abroad and shared without metadata may escape detection, and the implied truth effect suggests that partial labelling can lend credibility to unlabelled fakes.[87] Mislabelled

[84] Modern Dental Coll. & Research Ctr. v. State of Madhya Pradesh, (2016) 7 S.C.C. 353 (India); K.S. Puttaswamy v. Union of India, (2017) 10 S.C.C. 1 (India).
[85] K.S. Puttaswamy v. Union of India, (2019) 1 S.C.C. 1 (India).
[86] Anuradha Bhasin v. Union of India, (2020) 3 S.C.C. 637 (India).
[87] Pennycook et al., *supra* note 26.

authentic content may also lose credibility.[88] Rational connection is a low threshold and is met. For metadata and identifiers, the connection to attributing content to its originating tool is strong. The connection to traceability across platforms is weak, because any platform that strips metadata breaks the chain without breaching any duty imposed on the originating intermediary.[89]

3. **Necessity:** Following Chandra's caution, the alternatives must be stated concretely.[90] Three stand out. First, a context-sensitive label that is mandatory for realistic depictions of identifiable real persons or real-world events presented as factual, with a lighter, creator-controlled disclosure for evident satire, parody and fiction, would capture the cases that connect to Article 19(2) grounds while sparing expression that deceives no one. The European Union adopts essentially this approach.[91] Second, a duty on platforms to detect and display provenance data where it exists, together with targeted verification on risk signals, would serve the aim with less intrusion than universal pre-publication verification. California follows this model.[92] Third, graduated labels that reflect uncertainty, such as "possibly AI-generated", would reduce the harm of false positives. On the moderate test, the first alternative would realise the anti-deception aim in a real and substantial manner, which suggests that the undifferentiated scope of Rule 3(3)(a)(ii) and the universal character of Rule 4(1A) are not necessary in their present form.

---

[88] Altay & Gilardi, *supra* note 27.
[89] Mishra et al., *supra* note 31; C2PA IMPLEMENTATION GUIDANCE, *supra* note 29.
[90] Chandra, *supra* note 32.
[91] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), art. 50(4), 2024 O.J. (L 1689) 1.
[92] California AI Transparency Act, S.B. 942, 2023-2024 Leg., Reg. Sess. (Cal. 2024), *amended by* A.B. 853, 2025-2026 Leg., Reg. Sess., ch. 674 (Cal. 2025).

4. **Balancing:** The benefits of labelling realistic depictions of real people and events are significant, if empirically uncertain. The costs fall mainly on satirists, artists and political commentators, and on authentic recordings that automated verification wrongly labels, a form of misinformation the rules themselves would produce. The balance favours the core of the scheme and disfavours its margins.

### E. Article 14: Precision, Vagueness and Classification

*Shreya Singhal* struck down section 66A because terms such as "grossly offensive" and "menacing" gave no ascertainable standard and would chill protected speech.[93] The Bombay High Court applied similar reasoning to "fake or false or misleading" in *Kunal Kamra*.[94] The SGI provisions are more precise than either, but several terms invite the same concern. "Appears to be real, authentic or true" and "likely to be perceived as indistinguishable" do not specify the observer. "Materially alter" and "good-faith" in the exclusions call for judgment. "Computer resource of the intermediary" may mean a model, a server, an application or an account.[95] And "permanent" metadata is not technically attainable, since any embedded data can be removed by re-encoding; the feasibility qualifier saves the provision from commanding the impossible but leaves compliance uncertain.

Vagueness matters less in rules addressed to commercial actors than in criminal offences, and "prominent" is a familiar regulatory standard. The more serious Article 14 problem is that vagueness here operates through an intermediary that will resolve doubt by over-labelling. On classification, the distinction between SSMIs and other intermediaries' rests on scale and is rational. Excluding text is a defensible judgment about the distinct deceptive power of realistic audio-visual material. The absence of any size threshold in Rule 3(3) means that a small Indian start-up offering an image

[93] Shreya Singhal, (2015) 5 S.C.C. 1.
[94] Kunal Kamra, 2024 LiveLaw (Bom) 482.
[95] Mishra et al., *supra* note 31.

tool bears the same metadata and anti-tampering duties as a global platform. That is not manifestly arbitrary, but it bears on Article 19(1)(g).

### F. Article 21: Privacy, Anonymity and Informational Autonomy

*Puttaswamy* recognised informational privacy as part of Article 21 and required any intrusion to satisfy legality, a legitimate aim and proportionality, with procedural safeguards against abuse.[96] The provenance duty, on its text, identifies the intermediary's computer resource, not the user, so its facial privacy impact is limited.

The risk lies in linkage. An intermediary that issues a unique identifier for each output can match it against account logs, and Rule 3(1) (ca)(ii)(III) contemplates disclosing a violating user's identity to a complainant "in accordance with applicable law".[97] If intermediaries implement provenance through C2PA manifests carrying creator identity assertions, which the standard supports, metadata could disclose personal information to anyone who downloads the file. The rules neither require nor forbid this. The Digital Personal Data Protection Act, 2023 governs the intermediary's processing of personal data, but it does not address what provenance records may contain or how long identifier-to-account mappings may be retained.[98]

Anonymous expression has deep roots in democratic practice, as the United States Supreme Court recognised in protecting anonymous political leaflets in *McIntyre v. Ohio Elections Commission*.[99] Indian courts have not declared a free-standing right to anonymous speech, but *Puttaswamy*'s account of informational self-determination supports it. The Article 21 concerns are therefore real but contingent: they arise from implementation choices the rules leave open, not from the text itself. A rule that confined identifiers to system information and restricted linkage to legal process would remove most of them.

---

[96] Puttaswamy, (2017) 10 S.C.C. 1.
[97] IT Rules, 2021, r. 3(1)(ca)(ii)(III).
[98] The Digital Personal Data Protection Act, 2023, No. 22, Acts of Parliament, 2023 (India).
[99] McIntyre v. Ohio Elections Comm'n, 514 U.S. 334 (1995).

### G. Article 19(1)(g): Compliance Burdens

In *Internet and Mobile Association of India v. Reserve Bank of India*, the Supreme Court applied proportionality to a regulatory measure that cut off a business sector from banking services and found it disproportionate.[100] The SGI rules impose costs of a lesser order: detection systems, metadata pipelines, anti-tampering controls and, for SSMIs, pre-publication verification.

Regulatory cost is not itself a constitutional injury, and Article 19(6) permits reasonable restrictions in the public interest. The difficulty is that uncertainty multiplies cost. An intermediary that does not know which standard satisfies "other appropriate technical provenance mechanisms", or what "permanent" means, must over-engineer or accept legal risk. A published standard with a compliance safe harbour would cut that cost.

### H. Enforcement, Procedural Fairness and Chilling Effects

Rule 4(1A) relies on automated verification before publication. Detection tools err in both directions, and their accuracy across languages, compression levels and formats varies.[101] The rules provide no notice to a user whose declaration is overridden, no reasons and no tailored appeal against a label. A user may invoke the general grievance mechanism and, after the 2022 amendment, appeal to a Grievance Appellate Committee, but neither was designed for rapid correction of labels on time-sensitive content.[102]

The incentive structure points toward over-compliance. The deemed-failure proviso in Rule 4(1A), the loss of immunity under Rule 7 and the compressed timelines all penalise under-labelling and removal failures, while nothing penalises over-labelling or wrongful removal. This is a reasoned risk, not a measured effect, but it is the

[100] Internet & Mobile Ass'n of India v. Reserve Bank of India, (2020) 10 S.C.C. 274 (India).
[101] NIST AI 100-4, *supra* note 28.
[102] The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2022, G.S.R. 794(E), Gazette of India, Extraordinary, pt. II sec. 3(i) (Oct. 28, 2022) (India).

dynamic *Shreya Singhal* identified when it refused to make intermediaries judges of legality.[103] *Anuradha Bhasin*'s insistence on publication of restrictive orders points to a further remedy: transparency reporting on labels applied, declarations overridden and errors corrected.[104]

### I. How Far the Precedents Travel

The ratio of *Shreya Singhal* concerns vague speech offences and the conditions for intermediary takedown, not disclosure mandates. *Puttaswamy* and the Aadhaar judgment set the tests for privacy intrusions, but in a statutory identity scheme quite unlike content provenance. *Kunal Kamra* is the closest analogy, but it is a High Court decision under appeal, and its reasoning targeted a state unit empowered to declare truth, whereas the SGI label asserts nothing about truth.

Its observations on the State as arbiter of truth are persuasive, not binding, outside Maharashtra. Read together, the precedents support the analytical structure used here, but no decision resolves whether a factual label on lawful speech is a "restriction" within Article 19(2). That question remains open.

## IX. COMPARATIVE ANALYSIS

1. **European Union:** Article 50 of the AI Act divides duties between actors. Providers of generative systems must mark outputs in a machine-readable format detectable as artificially generated, as far as technically feasible. Deployers who generate or manipulate deepfakes must disclose that fact, but where the content forms part of an evidently artistic, creative, satirical or fictional work, the duty is limited to disclosure in a manner that does not hamper display or enjoyment of the work.[105] Article 50 has applied since 2 August 2026; the Digital Omnibus on AI gave systems already on the market

[103] Shreya Singhal, (2015) 5 S.C.C. 1; *see also* Arun, *supra* note 5.
[104] Anuradha Bhasin, (2020) 3 S.C.C. 637.
[105] AI Act, art. 50(2), (4).

until 2 December 2026 to implement machine-readable marking.[106] A voluntary Code of Practice published on 10 June 2026 sets out how providers and deployers may meet these duties.[107] Separately, the Digital Services Act lists prominent marking of deepfakes among risk-mitigation measures for very large online platforms.[108] India can learn from three features: separate machine marking from human disclosure, an express accommodation for art and satire, and a consulted code that gives technical content to a general duty.

2. **China:** China's Measures for Labeling AI-Generated Synthetic Content, effective 1 September 2025, require explicit labels perceptible to users and implicit labels embedded in metadata, supported by a mandatory national standard, and require users to declare synthetic content when publishing.[109] Distribution platforms that find no implicit label but detect other signs of generation apply a label indicating that content is possibly or suspectedly AI-generated.[110] India's design is closest to this model. The constitutional settings differ fundamentally, so the comparison yields only a design lesson: graduated labels acknowledge uncertainty in detection instead of asserting a conclusion the platform cannot verify.

3. **California:** The California AI Transparency Act, as amended in 2025, requires large generative AI providers to embed latent disclosures in image, video and

---

[106] *EU AI Act Article 50 Transparency Obligations Go Live*, BAKER BOTTS L.L.P. (Sept. 2026), https://www.bakerbotts.com/thought-leadership/publications/2026/september/eu-ai-act-article-50-transparency-obligations-go-live.

[107] Press Release, European Commission, Commission Publishes Code of Practice on Marking and Labelling AI-Generated Content (June 10, 2026), https://ec.europa.eu/commission/presscorner/detail/en/ip_26_1328.

[108] Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services (Digital Services Act), art. 35(1)(k), 2022 O.J. (L 277) 1.

[109] Cyberspace Admin. of China et al., Measures for Labeling Artificial Intelligence-Generated Synthetic Content (issued Mar. 2025, effective Sept. 1, 2025) (China); *China's AI-Labeling Measures and Mandatory National Standards Take Effect September 1*, LOEB & LOEB LLP (Mar. 2025), https://www.loeb.com/en/insights/publications/2025/03/chinas-ai-labeling-measures-and-mandatory-national-standards-take-effect-september-1.

[110] *China Releases New Labeling Requirements for AI-Generated Content*, INSIDE PRIVACY, COVINGTON & BURLING LLP (Mar. 2025), https://insideprivacy.com/international/china/china-releases-new-labeling-requirements-for-ai-generated-content.

audio output to the extent technically feasible, to offer users the option of a visible disclosure and to provide a free detection tool, with duties operative from 2 August 2026.[111] From 2027, large online platforms must detect compliant provenance data and show users whether it is available, rather than verify every upload. The statute also distinguishes "system provenance data" from "personal provenance data", which contains personal information.[112] That distinction offers India a precise tool for meeting the Article 21 concerns identified in Part 8.

**Table 1: Selected features of synthetic-content transparency regimes**

| Feature | India (IT Rules, as amended) | European Union (AI Act art. 50) | China (2025 Measures) | California (AI Transparency Act) |
|---|---|---|---|---|
| Human-facing label | Mandatory for all lawful SGI | Deployers disclose deepfakes; lighter duty for evident art and satire | Mandatory explicit label | Visible disclosure offered as an option |
| Machine-readable marking | Permanent metadata and unique identifier, to the extent | Providers mark outputs, as far as technically feasible | Mandatory implicit label in metadata | Latent disclosure, to the extent technically feasible |

[111] California AI Transparency Act; *New California AI Disclosure Rules Become Operative*, MORGAN LEWIS (Aug. 3, 2026), https://morganlewis.com/pubs/2026/08/new-california-ai-disclosure-rules-become-operative.
[112] California AI Transparency Act.

| Feature | India (IT Rules, as amended) | European Union (AI Act art. 50) | China (2025 Measures) | California (AI Transparency Act) |
|---|---|---|---|---|
| | technically feasible | | | |
| Platform role | SSMIs collect declarations and verify before publication | Risk mitigation for very large platforms under the DSA | Platforms check metadata; graduated labels | Large platforms detect and surface provenance from 2027 |
| Privacy safeguard in provenance | None specified | General data protection law | Not specified | Distinguishes system and personal provenance data |

## X. FINDINGS OF THE STUDY

1. The operative law is G.S.R. 120(E), in force from 20 February 2026. It defines SGI narrowly by reference to realism and indistinguishability, excludes text and three categories of routine use, and contains no exclusion for satire, parody or fiction. The continuous-label proposal and draft Rule 3(4) are not law, and the October 2026 advisory does not amend the rules.
2. The labelling and metadata duties are probably within section 79(2)(c) read with section 87(2) (zg). The prevention duty in Rule 3(3)(a)(i)(IV) and the deemed-failure proviso in Rule 4(1A) is more open to challenge as exceeding the Act and as inconsistent with *Shreya Singhal*.
3. A factual label engages Article 19(1)(a) but carries less weight than a prohibition. Its nexus with defamation, public order and incitement is close for

realistic depictions of real people and events presented as factual and remote for evident fiction and satire. Whether a disclosure mandate is a "restriction" under Article 19(2) is unresolved in Indian law.

4. The scheme passes the legitimate aim and rational connection stages. Its undifferentiated scope and universal pre-publication verification fail the necessity stage because context-sensitive labels and risk-based verification would achieve the aim in a real and substantial manner with less burden.

5. The provenance identifier is, on its text, attribution to a system rather than identification of a person. The Article 21 risk arises from linkage, retention and implementation choices that the rules do not regulate.

6. Terms such as "computer resource", "permanent" and "likely to be perceived as indistinguishable" are imprecise; the asymmetric feasibility qualifier, the absence of a size threshold in Rule 3(3) and the absence of a published technical standard raise compliance costs and encourage over-labelling. Metadata cannot deliver cross-platform traceability, though it supports attribution to the originating tool.

The framework lacks notice, reasons and appeal for users affected by labels, and nothing in it penalises over-labelling or wrongful removal.

## XI. SUGGESTIONS AND RECOMMENDATIONS

1. **Define the perception standard:** Rule 2(1) (wa) should refer to an ordinary reasonable viewer encountering the content in the context in which it is presented. This would anchor the test in a familiar legal standard and allow context, such as a satire channel or an obvious animation, to count.

2. **Adopt context-sensitive labels:** The mandatory prominent label should apply to realistic depictions of identifiable real persons and real-world events presented as factual. For evident satire, parody, fiction and art, the rules should accept a creator-chosen disclosure that does not hamper the work, following Article 50(4) of the AI Act. The continuous-label proposal should not be

adopted without evidence that persistent overlays improve recognition enough to justify their burden on short-form video.

3. **Make labels meaningful:** Labels should use plain language (made or altered with AI) linked to an explanation, and should distinguish confirmed from possible synthetic content. This responds directly to the evidence that audiences misread "AI-generated" as full automation.
4. **Protect privacy by design:** The rules should define "computer resource", state that unique identifiers must contain only system information, prohibit the embedding of personal data in provenance records without the user's choice, and limit identifier-to-account linkage to disclosure under legal process, with retention limits aligned to the DPDP Act. California's distinction between system and personal provenance data offers a workable template.
5. **Publish a technical standard:** MeitY should notify, after public consultation, an interoperable provenance standard compatible with open specifications such as C2PA, replace "permanent" with "durable and tamper-evident", extend the feasibility qualifier to visible labels and provide that compliance with the standard satisfies the rule.
6. **Recast verification as risk-based:** Rule 4(1A) should require verification where provenance signals, user reports or risk indicators warrant it, rather than across every upload, and should require SSMIs to display provenance data where it exists. The deemed-failure proviso should apply only on actual knowledge.
7. **Add procedural safeguards:** A user whose declaration is overridden or whose content is labelled should receive notice and reasons and should be able to seek review within a short, fixed period. SSMIs should publish periodic reports on labels applied, declarations overridden and errors corrected.
8. **Clarify consequences and legal basis:** Rule 3(1)(cb) should state that identity disclosure applies only to unlawful SGI under clause (a)(i), not to labelling failures. Substantive content obligations of this kind should move into primary

legislation, and draft Rule 3(4) should not be adopted unless every advisory given binding effect is published and open to challenge.

## XII. CONCLUSION

The central question was whether India's mandatory labelling and provenance obligations for synthetic content are constitutionally justified, precise and proportionate. The answer is mixed, and the mix matters. On scope, the operative obligations are those in G.S.R. 120(E), not the drafts or advisories. On statutory authority, the labelling and metadata duties rest on a plausible reading of section 79(2)(c), while the prevention and deemed-failure provisions are vulnerable. On Article 19, a factual label on realistic depictions of real people and events is a modest, justifiable burden tied to recognised grounds, but its extension to evident satire and fiction strains the connection.

On proportionality, the scheme has a legitimate aim and a rational connection but fails necessity at its margins. On Article 21, the identifier attributes content to systems, not persons; the risk lies in linkage. On precision and feasibility, undefined terms, the asymmetric feasibility qualifier and the absence of a published standard generate uncertainty that pushes intermediaries toward over-labelling. On safeguards, the reforms in Part 11 would preserve the aim at lower constitutional cost.

The strongest argument for the framework is that labelling allows realistic synthetic content to circulate lawfully instead of being blocked as deceptive, and that it serves audiences' right to know what they are seeing. The strongest objection is that the rules apply that disclosure, and pre-publication verification, without regard to whether any audience could be deceived, through intermediaries whose incentives all run toward excess. The findings on the existence and scope of the obligations reflect settled text; the conclusions on vires and proportionality are the author's assessment of open questions that no court has yet decided.

The balance would shift against the framework if the continuous-label proposal were finalised, if advisories extending labels to "misleading" content acquired binding force, or if evidence emerged of widespread mislabelling of authentic content. It

would shift in the framework's favour if MeitY adopted context-sensitive labels, a published standard and procedural remedies. The regime's constitutional future depends less on whether it labels than on how precisely it decides what to label.

## XIII. BIBLIOGRAPHY

### A. Legislation and Statutory Instruments

1. The Bharatiya Nyaya Sanhita, 2023, No. 45, Acts of Parliament, 2023 (India).
2. California AI Transparency Act, S.B. 942, 2023-2024 Leg., Reg. Sess. (Cal. 2024), *amended by* A.B. 853, 2025-2026 Leg., Reg. Sess., ch. 674 (Cal. 2025).
3. Cyberspace Admin. of China et al., Measures for Labeling Artificial Intelligence-Generated Synthetic Content (issued Mar. 2025, effective Sept. 1, 2025) (China).
4. The Digital Personal Data Protection Act, 2023, No. 22, Acts of Parliament, 2023 (India).
5. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2022, G.S.R. 794(E), Gazette of India, Extraordinary, pt. II sec. 3(i) (Oct. 28, 2022) (India).
6. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Amendment Rules, 2026, G.S.R. 120(E), Gazette of India, Extraordinary, pt. II sec. 3(i) (Feb. 10, 2026) (India).
7. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021, G.S.R. 139(E), Gazette of India, Extraordinary, pt. II sec. 3(i) (Feb. 25, 2021) (India).
8. The Information Technology Act, 2000, No. 21, Acts of Parliament, 2000 (India).
9. Regulation (EU) 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a Single Market for Digital Services (Digital Services Act), 2022 O.J. (L 277) 1.

10. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act), 2024 O.J. (L 1689) 1.

**B. Judicial Decisions**

1. Anil Kapoor v. Simply Life India, CS(COMM) No. 652 of 2023 (Del. H.C. Sept. 20, 2023) (India).
2. Anuradha Bhasin v. Union of India, (2020) 3 S.C.C. 637 (India).
3. Bijoe Emmanuel v. State of Kerala, (1986) 3 S.C.C. 615 (India).
4. Indian Express Newspapers (Bombay) (P) Ltd. v. Union of India, (1985) 1 S.C.C. 641 (India).
5. Internet & Mobile Ass'n of India v. Reserve Bank of India, (2020) 10 S.C.C. 274 (India).
6. K.S. Puttaswamy v. Union of India, (2017) 10 S.C.C. 1 (India).
7. K.S. Puttaswamy v. Union of India, (2019) 1 S.C.C. 1 (India).
8. Kaushal Kishor v. State of Uttar Pradesh, (2023) 4 S.C.C. 1 (India).
9. Kunal Kamra v. Union of India, W.P. (L) No. 9792 of 2023, 2024 LiveLaw (Bom) 482 (Bom. H.C. Sept. 26, 2024) (India).
10. McIntyre v. Ohio Elections Comm'n, 514 U.S. 334 (1995).
11. Modern Dental Coll. & Research Ctr. v. State of Madhya Pradesh, (2016) 7 S.C.C. 353 (India).
12. Nat'l Inst. of Family & Life Advocates v. Becerra, 585 U.S. 755 (2018).
13. Sec'y, Ministry of Info. & Broad. v. Cricket Ass'n of Bengal, (1995) 2 S.C.C. 161 (India).
14. Shayara Bano v. Union of India, (2017) 9 S.C.C. 1 (India).
15. Shreya Singhal v. Union of India, (2015) 5 S.C.C. 1 (India).
16. State of Tamil Nadu v. P. Krishnamurthy, (2006) 4 S.C.C. 517 (India).

17. Union of India v. Ass'n for Democratic Reforms, (2002) 5 S.C.C. 294 (India).

18. Union of India v. Motion Picture Ass'n, (1999) 6 S.C.C. 150 (India).

19. X Corp. v. Union of India, W.P. No. 7405 of 2025 (Kar. H.C. Sept. 24, 2025) (India).

20. Zauderer v. Office of Disciplinary Counsel, 471 U.S. 626 (1985).

**C. Books**

1. GAUTAM BHATIA, OFFEND, SHOCK, OR DISTURB: FREE SPEECH UNDER THE INDIAN CONSTITUTION (Oxford Univ. Press 2016).

**D. Journal Articles**

1. Aparna Chandra, *Proportionality in India: A Bridge to Nowhere?* 3 U. OXFORD HUM. RTS. HUB J. 55 (2020).

2. Chinmayi Arun, *Gatekeeper Liability and Article 19(1)(a) of the Constitution of India*, 7 NUJS L. REV. 73 (2014).

3. Gordon Pennycook et al., *The Implied Truth Effect: Attaching Warnings to a Subset of Fake News Headlines Increases Perceived Accuracy of Headlines Without Warnings*, 66 MGMT. SCI. 4944 (2020).

4. Jack M. Balkin, *Free Speech Is a Triangle*, 118 COLUM. L. REV. 2011 (2018).

5. Kate Klonick, *The New Governors: The People, Rules, and Processes Governing Online Speech*, 131 HARV. L. REV. 1598 (2018).

6. Robert Chesney & Danielle Keats Citron, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 CALIF. L. REV. 1753 (2019).

7. Sacha Altay & Fabrizio Gilardi, *People Are Skeptical of Headlines Labeled as AI-Generated, Even if True or Human-Made, Because They Assume Full AI Automation*, 3 PNAS NEXUS pgae403 (2024).

### E. Government and Institutional Documents

1. Press Release, European Commission, Commission Publishes Code of Practice on Marking and Labelling AI-Generated Content (June 10, 2026), https://ec.europa.eu/commission/presscorner/detail/en/ip_26_1328.

2. Ministry of Elec. & Info. Tech., Draft Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Second Amendment Rules, 2026 (Mar. 30, 2026) (India), https://www.meity.gov.in/static/uploads/2026/03/30591fc6e322dcbcc9dae84a0f02e9e7.pdf.

3. Ministry of Elec. & Info. Tech., Extension of Time for Inviting Feedback/Comments of Stakeholders on the Draft Amendments to the IT Rules, 2021, Including Additional Amendments in Rule 3(3)(a)(ii) (Apr. 21, 2026) (India), https://www.meity.gov.in/static/uploads/2026/04/ec197f1206279efb4964965f0dede6c1.pdf.

4. MINISTRY OF ELEC. & INFO. TECH., FREQUENTLY ASKED QUESTIONS ON THE INFORMATION TECHNOLOGY (INTERMEDIARY GUIDELINES AND DIGITAL MEDIA ETHICS CODE) AMENDMENT RULES, 2026 (2026), https://www.meity.gov.in/static/uploads/2025/10/065b6deb585441b5ccdf8be42502a49c.pdf.

5. Prabhat Katyayan Mishra et al., *Feedback on the Draft Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Second Amendment Rules, 2026*, CTR. FOR THE STUDY OF L. & GOVERNANCE, JAWAHARLAL NEHRU UNIV. (May 7, 2026), https://www.jnu.ac.in/sites/default/files/cslg/Feedback_ITGuidelines_DigitalMediaEthicsCode.pdf.

© 2026. LawFoyer International Journal of Doctrinal Legal Research (ISSN: 2583-7753)

### F. Technical Standards and Publications

1. COAL. FOR CONTENT PROVENANCE & AUTHENTICITY, C2PA IMPLEMENTATION GUIDANCE, https://c2pa.org/public-draft/Guidance.html (last visited Oct. 25, 2026).
2. COAL. FOR CONTENT PROVENANCE & AUTHENTICITY, C2PA TECHNICAL SPECIFICATION VERSION 2.2: SOFT BINDING RESOLUTION API (2025), https://c2pa.org/specifications/specifications/2.2/softbinding/Decoupled.html.
3. NAT'L INST. OF STANDARDS & TECH., NIST AI 100-4, REDUCING RISKS POSED BY SYNTHETIC CONTENT: AN OVERVIEW OF TECHNICAL APPROACHES TO DIGITAL CONTENT TRANSPARENCY (2024), https://doi.org/10.6028/NIST.AI.100-4.

### G. Online Sources

1. *2026 Amendments to India's Information Technology Rules*, LAW.ASIA (Oct. 2026), https://law.asia/india-it-rules-2026-ai-synthetic-media-compliance/.
2. Anjali Tripathi, *Exclusive: Amid Protests, MeitY Widens Labelling for Social Media Beyond Synthetic Content*, MEDIANAMA (Oct. 8, 2026), https://www.medianama.com/2026/10/223-meity-labelling-social-media-synthetic-content/.
3. *China Releases New Labeling Requirements for AI-Generated Content*, INSIDE PRIVACY, COVINGTON & BURLING LLP (Mar. 2025), https://insideprivacy.com/international/china/china-releases-new-labeling-requirements-for-ai-generated-content.
4. *China's AI-Labeling Measures and Mandatory National Standards Take Effect September 1*, LOEB & LOEB LLP (Mar. 2025), https://www.loeb.com/en/insights/publications/2025/03/chinas-ai-labeling-measures-and-mandatory-national-standards-take-effect-september-1.

5. *EU AI Act Article 50 Transparency Obligations Go Live*, BAKER BOTTS L.L.P. (Sept. 2026), https://www.bakerbotts.com/thought-leadership/publications/2026/september/eu-ai-act-article-50-transparency-obligations-go-live.

6. Indumugi C. & Jhanvi Anam, *IT Intermediary Amendment Rules, 2026 Contradict Their Purpose*, INTERNET FREEDOM FOUND. (Feb. 11, 2026), https://internetfreedom.in/it-intermediary-amendment-rules-2026-contradict-their-purpose/.

7. *IT Rules 2026*, IPLEADERS BLOG (2026), https://blog.ipleaders.in/it-rules-2026/.

8. *IT Rules, 2021*, INTERNET FREEDOM FOUND., https://internetfreedom.in/tag/it-rules-2021/ (last visited Oct. 25, 2026).

9. *Kunal Kamra v. Union of India*, GLOBAL FREEDOM OF EXPRESSION, COLUMBIA UNIV., https://globalfreedomofexpression.columbia.edu/cases/kunal-kamra-v-union-of-india/ (last visited Oct. 25, 2026).

10. *New California AI Disclosure Rules Become Operative*, MORGAN LEWIS (Aug. 3, 2026), https://morganlewis.com/pubs/2026/08/new-california-ai-disclosure-rules-become-operative.